

Improving land change detection based on uncertain survey maps using fuzzy sets

Stefan Leyk · Niklaus E. Zimmermann

Received: 13 June 2005 / Accepted: 1 June 2006 / Published online: 19 October 2006
© Springer Science+Business Media B.V. 2006

Abstract In this paper we present a method for correcting inherent classification bias in historical survey maps with which subsequent land cover change analysis can be improved. We linked generalized linear modelling techniques for spatial uncertainty prediction to fuzzy set based operations. The predicted uncertainty information was used to compute fuzzy memberships of forest and non-forest classes at each location. These memberships were used to reclassify the original map based on decision rules, which take into consideration the differences in identification likelihood during the historical mapping. Since the forest area was underestimated in the original mapping, the process allows to correct this bias by favouring forest, especially where uncertainty was high. The analyses were performed in a cross-wise manner between two study areas in order to examine whether the bias correction algorithm would still hold in an independent test area. Our approach resulted in a significant improvement of

the original map as indicated by an increase of the Normalized Mutual Information from 0.26 and 0.36 to 0.38 and 0.45 for the cross-wise test against reference maps in Pontresina and St. Moritz, respectively. Consequently subsequent land cover change assessments could be considerably improved by reducing the deviations from a reference change by almost 50 percent. We concluded that the use of logistic regression techniques for uncertainty modelling based on topographic gradients and fuzzy set operations are useful tools for predictively reducing uncertainty in maps and land cover change models. The procedure allows to get more reliable area estimates of crisp classes and it improves the computation of the fuzzy areas of classes. The approach has limitations when the original map shows high initial accuracy.

Keywords Predictive uncertainty modelling · Fuzzy sets · Land cover change analysis · Classification bias · Correction of survey maps · Area estimation

S. Leyk (✉)
Department of Geography, University of Zurich,
Winterthurerstrasse 190, CH-8057 Zurich,
Switzerland
e-mail: leyks@bluewin.ch

N. E. Zimmermann
Land Use Dynamics, Swiss Federal Research Institute
WSL, Zuercherstrasse 111, CH-8903 Birmensdorf,
Switzerland

Introduction

The analysis of land use and land cover change plays a key role in modern landscape research and landscape ecology (Baker 1989). Different fields such as landscape ecology, geographical information science and remote sensing have

contributed considerably to this area (Lu et al. 2004; Coppin et al. 2004). Usually, the length of a considered analysis period is limited to the time span for which aerial photography is available. To consider longer time periods, historical maps can be used (Kienast 1993). Comparing these maps with contemporary maps requires an extensive evaluation since they usually contain considerable inherent uncertainty (Plewe 2002; Leyk et al. 2005). We found few attempts to investigate this uncertainty. Only recently, Leyk and Zimmermann (2004) presented a method for predictive uncertainty modelling in field-based survey maps. With this approach a spatial field of inherent uncertainty can be predicted and mapped in relation to hypothesised topographic predictors relevant to historic surveyors. Thus with increasing distance from the valley floor, with decreasing distance to boundaries and with increasing altitude and steepness, the mapping quality was found to decrease. This is related to the surveyors' difficulties in earlier times to move in mountainous terrain with heavy equipment as well as the reliability of the historical triangulation network. Although it is not possible to assess all aspects of uncertainty, such models allow us to evaluate the original map and to improve map-based area estimates and change detections.

When considering natural objects, which have multiple memberships and vague definitions, approaches of fuzzy set theory have been shown to be adequate techniques for evaluation and processing. There is an abundance of research dealing with fuzzy sets in landscape analyses based on geographical information systems (GIS) and remote sensing (Burrough 1989; Fisher 2000; Robinson 2003). Fuzzy sets are frequently used for assessing the classification accuracy based on traditional confusion matrix approaches (Gopal and Woodcock 1994; Binaghi et al. 1999; Jäger and Benz 2000). Matsakis et al. (2000) have presented an approach to evaluate fuzzy partitions in the field of satellite image classification using plausibilistic closure. The improvement of land use map comparisons through fuzzy agreement maps and hierarchical fuzzy pattern matching has been demonstrated by Power et al. (2001). Ahlqvist et al. (2003) demonstrated a unified

representation of fuzzy and rough geographical data. Cheng et al. (2001) derived the spatial extent of uncertainties of classified objects based on different fuzzy object models.

Additionally, several studies have attempted to improve area estimates from land cover classes using fuzzy sets. Lewis and Brown (2001) used a generalised area-based confusion matrix for exploring the accuracy of area estimates. Their approach cannot be used for uncertainty-related issues since they defined fuzzy memberships of each class from sub-pixel area proportions that are assumed to sum to unity. Woodcock and Gopal (2000) recognised that area estimates from fuzzy values, which define for each location the degree of uncertainty to be correctly classified, exceed unity for small memberships. Here the limitation is that they used a linguistic scale which does not allow to make full use of fuzzy sets. The fundamental idea that the area of fuzzy geographical entities is a function of the alpha cut, has been used by Fonte and Lodwick (2004) to develop advanced fuzzy area operators. One remaining limitation is the difficulty in interpreting such estimates when several classes have memberships of different meanings. Generally, such approaches are designed for areas covered by reference data. In order to use such approaches for other regions they should be connected to predictive modelling tasks and their success should be measurable.

In this paper we present an approach that overcomes most of these limitations. We aim at demonstrating how predictive models of uncertainty in a historical map of the 19th century can be linked to fuzzy set approaches in order to improve the map accuracy and historical forest area estimations. The combined approach allowed us to remove bias from historical maps, thus increasing the accuracy of land change detection. Fuzzy sets were derived from predicted spatially explicit uncertainty maps. A reclassification procedure is presented to differentiate identification likelihoods of different landscape classes. The analysis was developed within two study areas in the Swiss Alps that allowed us to evaluate the method in independent test areas.

Fuzzy sets and fuzzy classifications

Fuzzy sets

Here, we briefly introduce the main aspects of fuzzy sets and fuzzy classifications that can be found in Burrough and McDonnell (1998), Dubois and Prade (2000) and Robinson (2003). The concept of fuzzy sets introduced by Zadeh (1965) provides ways of dealing with natural geographic variability and complexity when considering vegetation or land cover classes. A membership function defining the degree of membership of an element to each of the existing classes allows us to address the gradual transition and multi-memberships in space. Thus by using fuzzy sets one can overcome the limitations of crisp sets in which only full or no membership can be assigned to each unit (Klir and Wierman 1999). A formal definition of a fuzzy set can be given as follows: Let $X = \{x\}$ be a finite set or space of geographical entities x of one considered class (universe of discourse). A fuzzy set A of this space is defined by a membership function μ_A in the ordered pairs $A = \{x, \mu_A(x)\}$ for each $x \in X$. The membership values range from zero (no membership) to one (full membership) on a continuous scale and with gradual transition: $\mu_A: X \rightarrow [0,1]$.

There are different approaches to assign membership functions appropriately resulting in linguistic (Power et al. 2001; Woodcock and Gopal 2000), categorical or continuous (Brown 1998; Andréfouët et al. 2000) membership values. Robinson (2003) describes the fundamental concerns of choosing the right membership function as standard function, problem-specific function or from data-driven approaches in GIS-related applications.

Fuzzy classification

Land cover maps often consist of classes that are continuous in nature. Thus many locations are expected to show gradual or multi-memberships to different classes, in particular close to boundaries. To account for these characteristics fuzzy classifications (Foody 1996; Bolliger et al. 2005) or fuzzy partitions (Matsakis et al. 2000) have been applied. They can be obtained by fuzzy classifiers, softening of crisp classes or neural networks (Woodcock and Gopal 2000).

A pixel (x_j) of the universe of discourse X ($X = \{x_j\}_{j=1,\dots,N}$) covers more than one land cover class (i) especially when the location is in the transition zone between two classes, which are neither well-defined nor spatially distinct. Therefore, a fuzzy classification is a family of M fuzzy sets $\{\mu_1, \dots, \mu_M\}$ such that $\forall i = 1, \dots, M, \mu_i \neq \emptyset, \mu_i \neq X$ (Dubois and Prade 2000). Given these conditions, a fuzzy partition into M classes of X can be considered as an M -tuple $\mu_{i,i=1,\dots,M}$ of functions from X into $[0,1]$ such that:

$$\forall j = 1, \dots, N, \exists i = 1, \dots, M \setminus \mu_{ij} > 0, \quad \text{and} \quad (1)$$

$$\forall i = 1, \dots, M, \exists j = 1, \dots, N \setminus \mu_{ij} > 0, \quad (2)$$

where μ_{ij} denotes the membership degree of pixel x_j of X in fuzzy class i . N and M denote two integers such that $N \geq M \geq 2$ and X denotes a set of N elements (universe of discourse) as described above. The membership μ_{ij} has been interpreted as possibility that element x_j of X belongs to class i (Krishnapuram and Keller 1993). Adding the condition $\forall j = 1, \dots, N, \sum_{i=1,\dots,M} \mu_{ij} = 1$ to Eqs. (1) and (2), probabilistic fuzzy partitions are obtained (Ruspini 1969; Bezdek 1981). In remote sensing, fuzzy partitions are frequently considered to be probability distributions where membership degrees nearly agree with the land cover proportion inside a considered pixel. Thus they sum to 1. In cases where they do not sum to 1, they cannot be considered probabilistic. In some cases such a partition is treated as possibilistic fuzzy partition (Matsakis et al. 2000).

Area estimation from fuzzy classes

Various authors described the area of a fuzzy class as a function of class membership. In cases where the fuzzy memberships are representing the uncertainty of belonging to a class they cannot be seen as proportion of that class within the pixel as is frequently done in remote sensing (Lewis and Brown 2001). Thus the area of fuzzy class A_i has to be represented as a function of the alpha cut level, which indicates at what level of fuzzy membership a pixel is considered to belong to class i exclusively. Given the condition $\forall i = 1, \dots, M, \exists x_j \setminus \mu_{ij} = 1$ which implies that

$h(\mu_i) = 1$, where h is the height of the fuzzy set (Klir and Wierman 1999), fuzzy class i can be considered as a normal fuzzy set. Thus it can be represented as a group of alpha cuts, for $\alpha \in [0,1]$. The area of a fuzzy object represented by a normal fuzzy set is the sum of the areas of the pixels of this object, which belong to the defined alpha cut level. Thus the areas of pixels j of class i , for which $\mu_{ij} \geq \alpha$, are summed to represent the area of the fuzzy object i as a decreasing left continuous function of the alpha cut level (Fonte and Lodwick 2004). For low alpha cut levels the area of the fuzzy class exceeds the area of a (crisp) reference class (Woodcock and Gopal 2000) resulting in proportions > 1 (Fig. 1).

One problem with this approach is that for different classes i the meanings of their memberships are assumed to be equal. In nature the classes have very different properties and different transition characteristics. To solve this problem, the respective alpha cut levels have to be computed in relation to all classes simultaneously. This is one prerequisite to achieve more reliable area estimates and defuzzification results if two or more classes are involved.

Material and methods

In the following, we describe a sequence of steps that were necessary to improve an original crisp

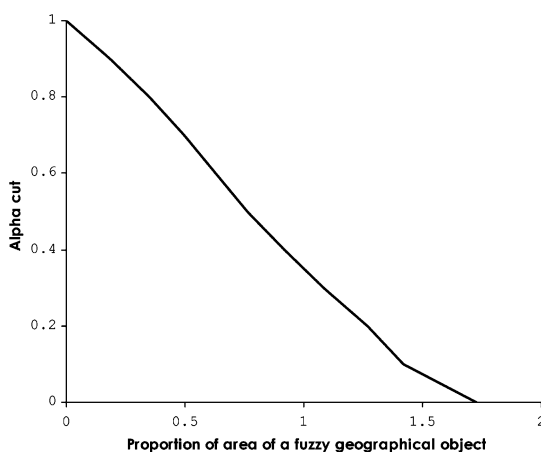


Fig. 1 Graphical representation of the proportion of the area of a fuzzy geographical entity that is mapped as a function of the alpha cut value. The example shows a forest patch from the study area St. Moritz

forest/non-forest map for more reliable area estimations based on a combination of predictive uncertainty modelling and bias correction using fuzzy set theory (Fig. 2). First, the predictive modelling procedure (Leyk and Zimmermann 2004) is presented, which allowed us to predict the spatial field of uncertainty in mapping forest as a function of topographic features. Then we demonstrate how the modelled uncertainty, which was converted to certainty, was used to transform the original crisp classes into two fuzzy classes, forest and non-forest, respectively, to account for transition zones and gradual memberships.

Finally, we present decision rules for the retransformation of the fuzzy classes into new crisp classes. This step aimed at improving classification accuracy and area estimations by correcting the bias in the mapped forest area, which was expected to be underestimated. The basic idea was that during the historical fieldwork, which was object-driven, different identification likelihoods have to be expected for forest and non-forest, respectively. Where “forest” was mapped, we could assume that forest had been found approximately at this position since the topographer had identified this object. Whenever forest could not be identified, the alternative class “non-forest” that was not extensively searched for, was assigned. Consequently, the memberships of forest and non-forest cannot be treated equally, especially in areas of low certainty. To account for

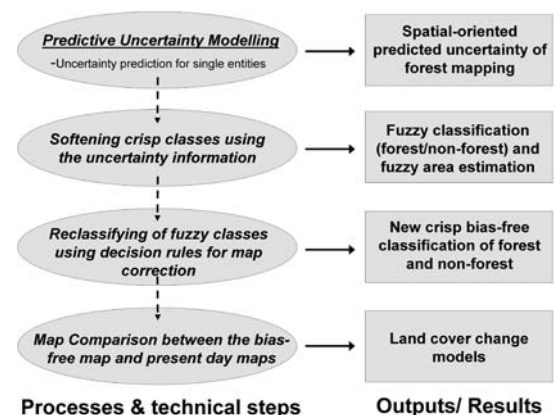


Fig. 2 Process flow of the analyses. Predictive uncertainty modelling, uncertainty-based fuzzy set generation and rule-based bias correction to produce improved forest/non-forest maps. These maps are applied for improved land change analysis by map comparison

this, dynamic weighting factors were examined to find the optimal retransformation, i.e. where the improvement in classification accuracy and thus the correction of bias reached their maximum.

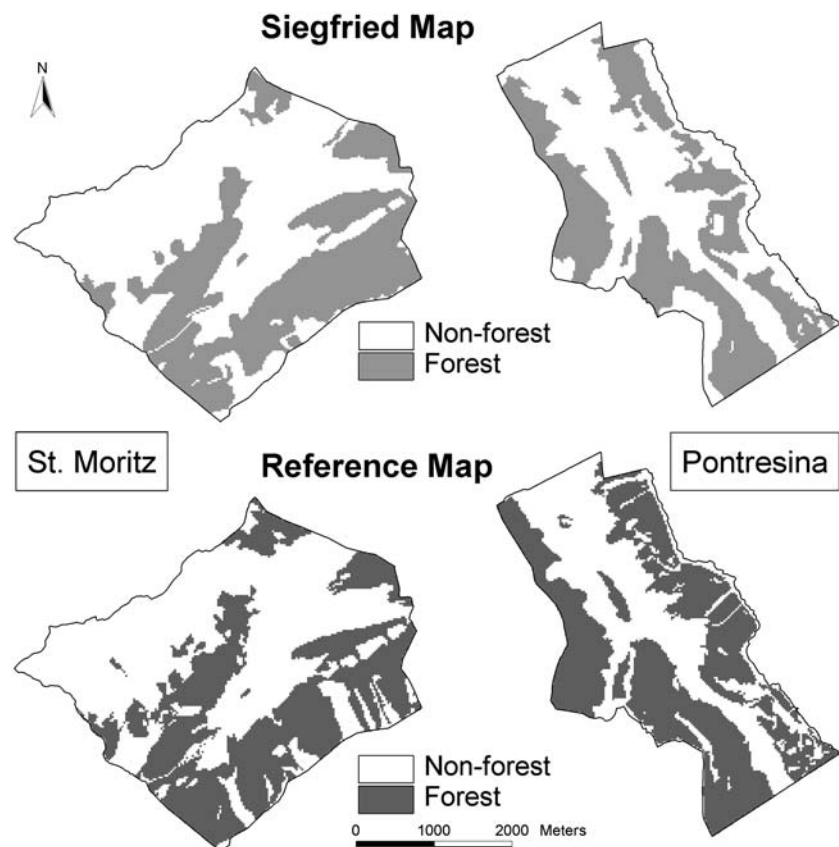
Study area, historical, and modern maps

The historical map to evaluate was the first edition of the so-called Siegfried Map series – the historical Swiss National topographic map published between 1870 and 1920. The special new feature here, compared to earlier maps, was the delineation of the forest cover. This allowed us to derive spatially explicit forest cover information for the entire area of Switzerland for that time period. Our study areas included the municipalities of Pontresina, where the uncertainty model had been developed (Leyk and Zimmermann 2004), and St. Moritz (Canton Grisons) (Fig. 3). For both communities we had access to accurate, local reference maps, which are otherwise scarce for the considered time period. This enabled us to

cross-wise calibrate and test the uncertainty model, and to assess its robustness and generality. These reference maps – communal maps at a scale of 1:5,000 or 1:10,000 with a high degree of detail – belong to a map series that preceded the official cadastral mapping in Switzerland at that time. For geo-rectification of all maps involved, we chose the projection of the present-day Swiss topographic map. There was evidence of some differences regarding the mapping of forest because avalanche tracks, larger bedrocks and forest gaps were rarely registered in the Siegfried Map while they were in the reference maps. As described in Leyk et al. (2005) there are various potential sources of inherent uncertainty in the Siegfried Map such as vague definitions, mapping errors or ambiguous concepts. The resulting spatial patterns of uncertainty have to be considered interrelated resulting in complex fields.

In an example of an application, we compared the improved historical map with a modern forest/non-forest map, which we derived from the

Fig. 3 Forest/Non-forest classes derived from Siegfried Map and reference maps of the two study areas Pontresina and St. Moritz



Swiss topographic map at a scale of 1:25,000 (LK25 maps from 2003, © Swisstopo). The map was available in digital form and allowed the distinction of the closed canopy forest layer.

A regional predictive uncertainty model

In an earlier study we presented a method for predictively modelling the spatial pattern of inherent mapping uncertainty E using topographic and distance measures as predictors (Leyk and Zimmermann 2004). The rationale for this approach was to mimic the difficulties of surveyors to identify forest and non-forest for mapping at the time of the Siegfried Map. By applying Generalised Linear Models (GLM) using the logit link function, we were able to incorporate these relationships into a predictive uncertainty model of mapping forest/non-forest classes. The general logistic regression model had the form (Eq. 3):

$$\text{logit} = \log(E/(1 - E)) = \alpha + X^T \beta, \quad (3)$$

where α is the regression intercept, X represents a vector of p predictor variables (X_1, \dots, X_p) of any power T , and β denotes a vector of p regression coefficients (β_1, \dots, β_p) which were determined for each predictor. Finally, E is the expected mean value of the response variable. GLMs perform regressions in a transformed space, overcoming the restrictions of ordinary least-squares models (e.g., E is bounded between 0 and 1).

In our recent study, the independent variables (X_1, \dots, X_p) distance to forest boundary, elevation difference to valley bottom and slope carried highest predictive power for local uncertainty E of a target pixel, as the model tests showed. We computed E as $1 - \kappa$, where κ is Kappa within local windows of 100×100 m. In each window the accuracy of the Siegfried Map was tested against the reference maps and E was recorded at the central pixel. We distributed the local windows systematically at regular distances larger than the window sizes to avoid overlaps and to reduce spatial autocorrelation and pseudo-replication. The layer of E was intersected with the independent variables to calibrate a predictive statistical model using GLMs in Splus

(Insightful 2001). The model allowed us to map E in a spatially explicit form, where $E = 0$ indicates a perfect fit and $E = 1$ means no agreement at all.

We applied standard stepwise regression procedures starting from a full model including all predictors in linear and quadratic form, and allowing for both backward and forward selection in order to optimize the models based on the Akaike Information Criterion (AIC). We reported AIC and the *adjusted* D^2 values for the goodness-of-fit of the models. AIC is a criterion for selecting among nested models by minimizing its value over choices of a number of parameters to form a trade-off between the fit of the model and its complexity. The *adjusted* D^2 is a measure of deviance reduction in the case of Maximum-Likelihood based estimation including the number of observations and the number of predictors (Guisan and Zimmermann 2000). Additionally we reported the final model parameters. Independent tests were performed by comparing the model calibrated at the respective other area with E as derived from the comparison with local communal maps. These tests were based on mean absolute error (*MAE*), which is the weighted average of the absolute errors, Spearman's rank correlation ρ , and the *G*-value that indicates the relative improvement of the model over the null model, i.e., the mean of E within the calibration data set. *MAE* ranges between zero (perfect fit) and infinity (no fit at all); ρ and *G* report high accuracy if values are close to 1 and random agreement if values are 0.

Generating fuzzy memberships from spatial fields of uncertainty

The predicted uncertainties E_j were used for defining fuzzy memberships of forest and non-forest by softening the original crisp classification of the Siegfried Map CL. For each pixel j $CL_j = 1$ if j was located inside the forest area and $CL_j = 0$ if j was located inside the non-forest area, respectively (Fig. 5a). It is important to recall that the uncertainty model does not account for a specific type of uncertainty. It simply provides information on the spatial fields of inherent uncertainty (Leyk et al. 2005) quantified by E (Eq. 3). We used the semantic import (SI)

(Robinson 1988) approach to assign the fuzzy membership functions hereafter. This was done by computing for each pixel j the degree of certainty C_j with $C_j = 1 - E_j = \kappa$ and importing C as external data to relax CL. Thus the degree to which pixel j belongs to forest and non-forest depends on C_j and CL_j .

The modelled values C_j showed local minima close to the vectorized forest boundary lines, meaning that distinguishing forest from non-forest is least certain near the mapped forest boundary. The values of C_j increase towards 1 in both directions away from the boundary (Figs. 4 and 5d). To derive fuzzy sets for forest (F) and non-forest (NF), C had to be transformed into two one-sided continuous surfaces based on the membership functions μ_F for F and μ_{NF} for NF. Inside a class the classification certainty is usually high. This certainty decreases when approaching the border of this class. Thus inside a forest ($CL = 1$), we can take $\mu_{F,j} = C_j$ to define F. When now moving into the non-forest area ($CL = 0$), the resulting memberships $\mu_{F,j}$ of F have to continue decreasing, tending towards zero with increasing distance from the forest boundary (Fig. 4). In general, this means that in order to compute the

memberships $\mu_{F,j}$ where $CL_j = 0$ and $\mu_{NF,j}$ where $CL_j = 1$ that is done the same way.

Using the example of $\mu_{F,j}$ where $CL_j = 0$, we need to know the mirroring value for each pixel outside the mapped forest area to convert the C_j values into $\mu_{F,j}$. If C_j at the forest boundary would always have the same value, the problem would be easily solved. But because the values C_j vary considerably along the forest boundary (due to topographic effects) and because this boundary has complex shapes, we need to (1) define the pixel values at the forest boundary, and (2) interpolate these pixel values from the forest boundary into the non-forest space. By this, each pixel j where $CL_j = 0$ has the respective mirror value associated (Fig. 4). For this reason, pixels C_j whose centres were less than half the size of a pixel away from the digitized vector boundary line in the Siegfried Map were selected to define $C_{\text{bound},j}$, the certainty values at the boundary. Inverse distance weighted interpolation (IDW) of C_{bound} has been applied to compute two surfaces, B_F representing the interpolated C_{bound} values where $CL = 0$, and B_{NF} representing the interpolated C_{bound} values where $CL = 1$ (Figs. 4 and 5b, c). IDW was based on the values of the two

$$\mu_{F,j}(C_j, B_{F,j}, CL_j) = \begin{cases} C_j & \text{if } CL_j = 1 \\ B_{F,j} - (C_j - B_{F,j}) & \text{if } CL_j = 0 \text{ and if } B_{F,j} > (C_j - B_{F,j}) \\ B_{F,j} & \text{if } CL_j = 0 \text{ and if } B_{F,j} > C_j \\ 0 & \text{if } CL_j = 0 \text{ and if } B_{F,j} < (C_j - B_{F,j}) \end{cases}, \quad (4)$$

$$\mu_{NF,j}(C_j, B_{NF,j}, CL_j) = \begin{cases} C_j & \text{if } CL_j = 0 \\ B_{NF,j} - (C_j - B_{NF,j}) & \text{if } CL_j = 1 \text{ and if } B_{NF,j} > (C_j - B_{NF,j}) \\ B_{NF,j} & \text{if } CL_j = 1 \text{ and if } B_{NF,j} > C_j \\ 0 & \text{if } CL_j = 1 \text{ and if } B_{NF,j} < (C_j - B_{NF,j}) \end{cases}. \quad (5)$$

values $\mu_{F,j}$ of the forest membership surface within the non-forest class ($CL = 0$), the C_j values need to be mirrored with respect to the C_j values that we find at the forest boundary. This transformation of C_j outside the respective original crisp class is a crucial point in our analysis. Hereafter, we describe our method to define the

nearest neighbours and a weight of 1.0. Thus B_F and B_{NF} represent the spatial fields of the mirror values based on C_{bound} (Fig. 4).

The surfaces CL , C , B_F and B_{NF} could now be used to compute μ_F and μ_{NF} for each pixel j : Let $X = \{x_j\}$ be the space of objects of one class covering one study area. The two fuzzy sets F and NF of X

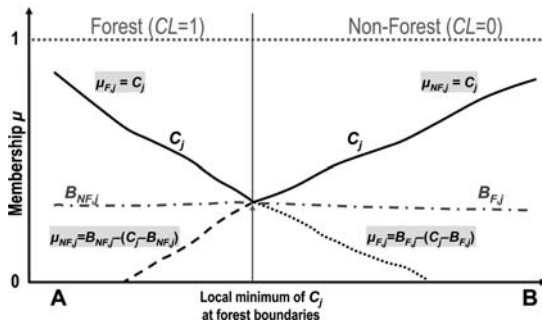


Fig. 4 Schematic illustration of the computation of fuzzy memberships of forest (μ_F) and non-forest (μ_{NF}) according to Eqs. 4 and 5 along a line transect between points A and B. Values μ_F inside the original forest area ($CL = 1$) and μ_{NF} inside the original non-forest area ($CL = 0$) are directly derived from certainty values C . Values μ_F where $CL = 0$ (dotted black line) and values μ_{NF} where $CL = 1$ (dashed black line) are mirrored against C using the interpolated surfaces B_F and B_{NF} (grey dashed and dotted line) to compute continuous values

were defined by μ_F and μ_{NF} in the ordered pairs $F = \{x_j, \mu_F(x_j)\}$ and $NF = \{x_j, \mu_{NF}(x_j)\}$, respectively, for each $x_j \in X, \mu_F: X \rightarrow [0,1], \mu_{NF}: X \rightarrow [0,1]$. The membership values $\mu_{F,j}$ and $\mu_{NF,j}$ at locations j , $j = 1, \dots, N$, were computed as presented in Eqs. 4 and 5, where C_j is the local certainty at location j , derived from $C = 1 - E$ (E is from Eq. 3), $B_{F,j}$ and $B_{NF,j}$ are the interpolated C_{bound} values at locations j , where $CL_j = 0$ and $CL_j = 1$, respectively, and CL_j indicates the Siegfried forest (1)/non-forest (0) classification (Fig. 5a). At locations j , where $CL_j = 1$, $\mu_{F,j}$ was defined as $\mu_{F,j} = C_j$ (Eq. 4), where $CL_j = 0$, $\mu_{NF,j}$ was defined as $\mu_{NF,j} = C_j$ (Eq. 5). This is a direct conversion of C_j into memberships.

The assignment of the remaining locations was more complex: At locations j , where $CL_j = 0$, $\mu_{F,j}$ of F was computed as $\mu_{F,j} = B_{F,j} - (C_j - B_{F,j}) = 2B_{F,j} - C_j$. By this for each pixel j , $\mu_{F,j}$ is computed as the mirror value of C_j based on $B_{F,j}$ (Fig. 4). The values of $\mu_{F,j}$ decrease with increasing C_j when moving away from the forest boundary within the non-forest area ($CL = 0$). Due to spatial resolution and interpolation effects C_j may be slightly smaller than $B_{F,j}$ in some cases close to the boundary. In such cases $\mu_{F,j}$ was defined as $\mu_{F,j} = B_{F,j}$ since $C_{\text{bound},j} \geq B_{F,j}$ can be assumed. Where $C_j > 2 \times B_{F,j}$, $\mu_{F,j}$ becomes negative. Since the increase of C is related to a decrease in μ_F , negative values indicate a forest membership of zero

(Eq. 4). This procedure was applied to compute $\mu_{NF,j}$ of NF for locations j , where $CL_j = 1$ (Eq. 5), accordingly. With this approach, which is illustrated in Fig. 5 by a matrix of computed values within a subset window, we derived two fuzzy sets, F and NF (Fig. 5e, f) on a continuous scale.

Area bias correction from fuzzy memberships

The fuzzy classification was used in a rule-based reclassification procedure to create a new crisp forest (1)/non-forest (0) map (CL_{Defuzz}) of improved accuracy by correcting bias. This was necessary to compare the historical map with contemporary land use maps for land change detection. Such a reclassification allowed us to assess the gain in accuracy of the historical maps after bias removal and to compute more reliable area estimates at the same time.

The decision rules have the effect that forest that is underestimated is weighted due to the lower identification likelihood during the historical mapping. For this weighting a new variable THR ($THR \in [0,1]$) is introduced. It represents the threshold value for C to differentiate between pixels j of high certainty ($C_j \geq THR$) and low certainty ($C_j < THR$). Where C_j was high, such as in valleys or close to roads forest boundaries are trustworthy and expected to be accurately delineated. At these locations μ_F and μ_{NF} could be compared directly to each other. These pixels are assigned to the class “forest” ($CL_{\text{Defuzz},j} = 1$) where $\mu_{F,j} \geq \mu_{NF,j}$ and to “non-forest” ($CL_{\text{Defuzz},j} = 0$) where $\mu_{F,j} < \mu_{NF,j}$. Where C_j was low, e.g., on steep slopes away from roads, pixels were assigned to $CL_{\text{Defuzz},j} = 0$ where $\mu_{F,j} = 0$ and to $CL_{\text{Defuzz},j} = 1$ where $\mu_{F,j} > 0$. This means that a pixel j was reclassified to forest ($CL_{\text{Defuzz},j} = 1$) even at very low values of μ_F if certainty C_j was low ($C_j < THR$). The decision rules that are based on μ_F , μ_{NF} , C and THR have the form (Eq. 6):

$$CL_{\text{Defuzz},j} = \begin{cases} 1 & \text{if } (C_j \geq THR \text{ and } \mu_{F,j} \geq \mu_{NF,j}) \\ 1 & \text{if } (C_j < THR \text{ and } \mu_{F,j} > 0) \\ 0 & \text{if } (C_j \geq THR \text{ and } \mu_{F,j} < \mu_{NF,j}) \\ 0 & \text{if } \mu_{F,j} = 0 \end{cases}, \quad (6)$$

Fig. 5 Derivation of fuzzy classes F for forest (e) and NF for non-forest (f) in Eqs. 4 and 5 using (a) the crisp original Siegfried classification CL , the interpolated surfaces (b) B_F and (c) B_{NF} and (d) the certainty values C . Below, numeric values from the sub-window in the grey scale images are presented as an example

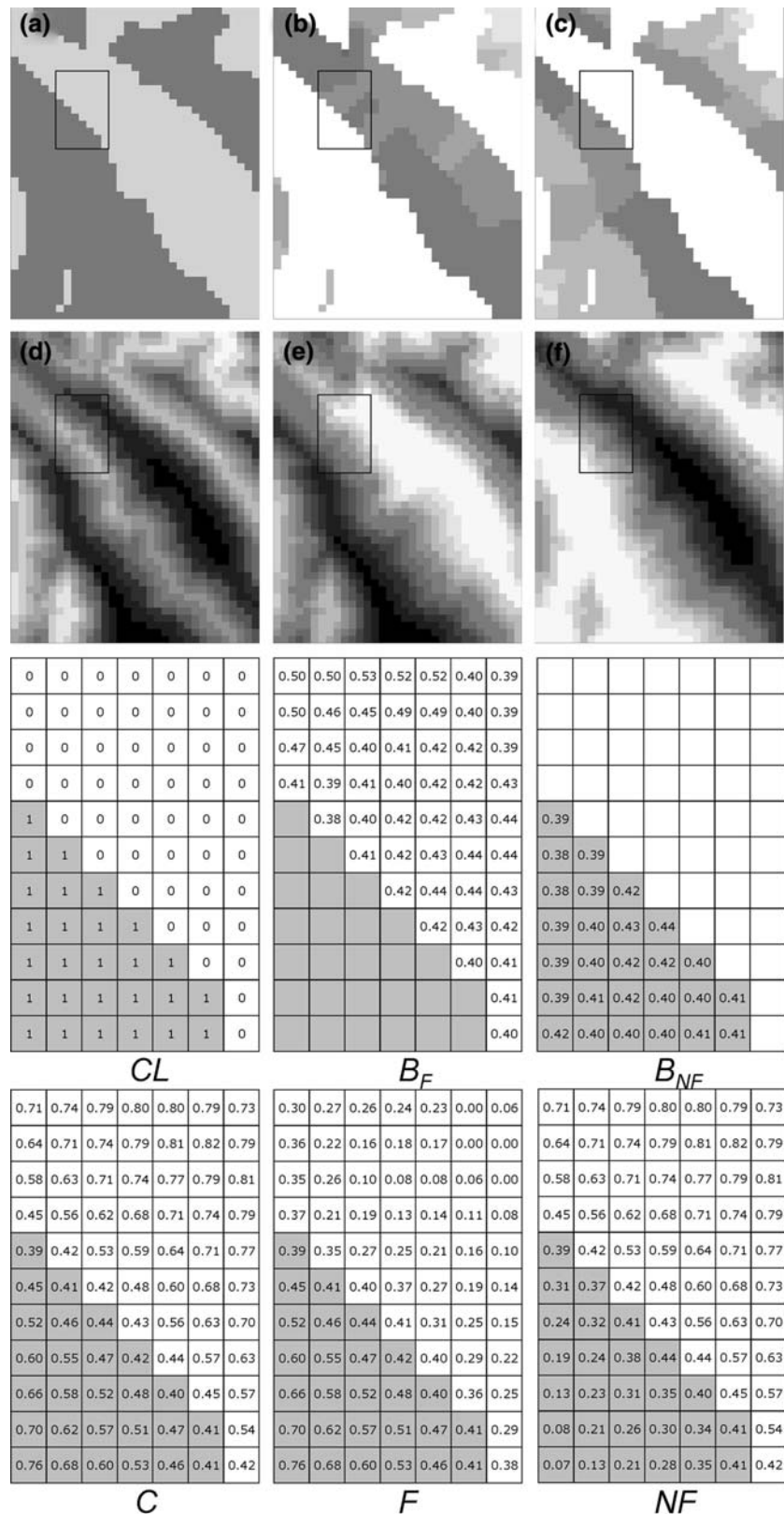
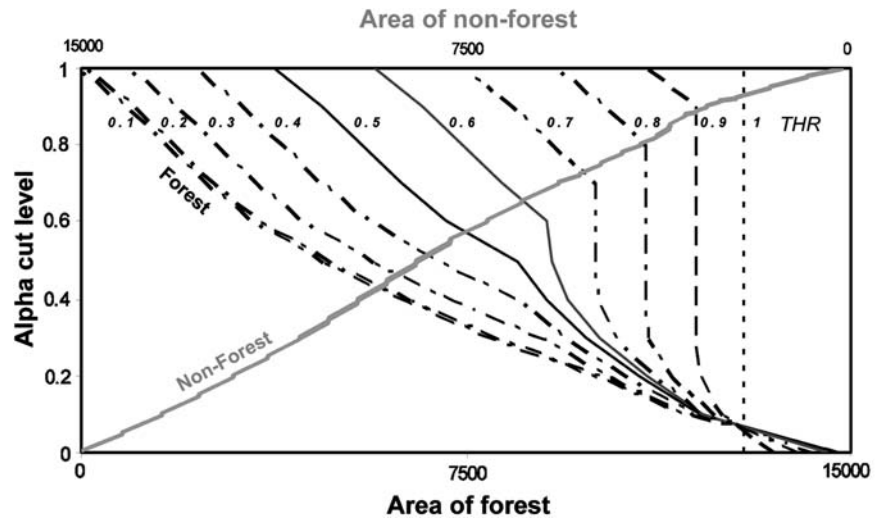


Fig. 6 Fuzzy areas (unitary pixels) of non-forest and forest. Non-forest areas, represented by the grey line oriented from the right to the left, are computed as function of the alpha level only. The fuzzy area of forest additionally considers THR which runs from 0 to 1, indicated by italic numbers close to the graphs



where $CL_{Defuzz,j}$ is the value of the new reclassified crisp set CL_{Defuzz} at location j . We iteratively computed all possible new crisp classifications for $THR \in [0,1]$ in steps of 0.01 and compared them with the reference maps for assessing the new accuracies. We computed three measures of accuracy: the percentage of correctly classified pixels (PCC), Cohen's Kappa (Cohen 1960), and the Normalized Mutual Information (NMI) (Forbes 1995). The best new classification was found where the gain in accuracy against the reference maps reached its maximum. For this value of THR the bias in forest area can be considered corrected. This procedure is flexible since it can be used to find out (1) which class (direction) is underestimated (i.e., increasing accuracy of CL_{Defuzz} when weighting the class) and (2) to what degree this class has to be weighted to

approach to new areas of similar topographic characteristics. This has to be validated through an extended analysis.

If the new map shows higher accuracy, it could be used for more reliable area estimation of the crisp classes, forest and non-forest. The decision rules (Eq. 6) can also be used to derive the areas of the fuzzy classes, F and NF , directly. The fuzzy area of NF is considered a function of the alpha-cut and was called $Area_{NF}(\alpha)$ (Fig. 6). The result was a family of crisp sets, where the areas of all pixels for which $\mu_{NF,j} \geq \alpha$ were summed to derive $Area_{NF}(\alpha)$. The fuzzy area of forest was computed as a family of crisp sets, each as a function of two independent variables, THR and α , and is called $Area_F(THR, \alpha)$. The equations used for computing $Area_{NF}(\alpha)$ and $Area_F(THR, \alpha)$ are given below (Eqs. 7 and 8):

$$Area_{NF}(\alpha) = \sum_{j=1}^n x_j, \text{ where } (\mu_{NF,j} \geq \alpha), \quad (7)$$

$$Area_F(THR, \alpha) = \sum_{j=1}^n x_j, \text{ where } (C_j \geq THR \text{ and } \mu_{F,j} \geq \alpha) \text{ or } (C_j < THR \text{ and } \mu_{F,j} > 0), \quad (8)$$

correct bias. For a validation of the predictive power of the analysis, the optimal values of THR had to be similar in both study areas. Such a similarity could indicate the applicability of the

where x_j are the areas of pixels j . $Area_F$ (Eq. 8) is computed for all combinations of $THR \in [0,1]$ and $\alpha \in [0,1]$ and presented for $THR = 0.1$ up to $THR=1$ (steps of 0.1) for illustration (Fig. 6). For

very small values of THR, Area_F equals a function of α only. For each THR the amplitude of the interval of the values Area_F for $\alpha \in [0,1]$ is determined by the number of pixels fulfilling the conditions $C_j \geq \text{THR}$ and $\mu_{F,j} \geq \alpha$, meaning that either $C_j = \mu_{F,j}$ if $\alpha \geq \text{THR}$ or $C_j \neq \mu_{F,j}$ if $\alpha < \text{THR}$. This amplitude decreases with increasing THR that is discernable by a shift of the graph to the right for $\alpha > 0.1$ and to the left for $\alpha \leq 0.1$ (Fig. 6). The main reason for the shift to the right is the increasing proportion of pixels fulfilling the conditions $C_j < \text{THR}$ and $\mu_{F,j} > 0$ that causes the correction of forest area (Eq. 6). Accordingly, steeper sections of the graphs for $\alpha < \text{THR}$ below a breakpoint for each THR at $\alpha = \text{THR}$ can be observed. These sections tend to be perpendicular for $\text{THR} \geq 0.7$ resulting in an entirely perpendicular graph for the extreme $\text{THR} = 1$. The shift to the left for $\alpha \leq 0.1$ is due to the lower amount of pixels fulfilling the condition $C_j \geq \text{THR}$ for large THR. For very large THR and very small α overestimations of forest area can thus be expected.

Results

Predictive uncertainty modelling and the derivation of fuzzy sets

The models calibrated independently for the two regions indicated reasonable predictive power for uncertainty E when tested in the respective other study area. The model calibrated in the St. Moritz area performed slightly

better. The step-wise regression resulted in a deviance reduction (*adjusted* D^2) of 35 percent in St. Moritz and 31 percent in Pontresina (Table 1), respectively. When compared with the independent test data, we received a high agreement and acceptable errors for the predicted values E ($\rho = 0.67$ and $\text{MAE} = 0.19$ when calibrated in St. Moritz and tested in Pontresina; $\rho = 0.69$ and $\text{MAE} = 0.16$ when calibrated and tested vice versa). The G -values indicated significant improvements over the null model ($G = 38\%$, $G = 44\%$, respectively, Table 1). The regression coefficients of both models have the same signs and are very similar values (Table 1), indicating a constant model performance over two test areas of different topographic characteristics. The cartographic representations of the distributions of certainty C and the resulting fuzzy sets F and NF for both study areas are shown in Fig. 7.

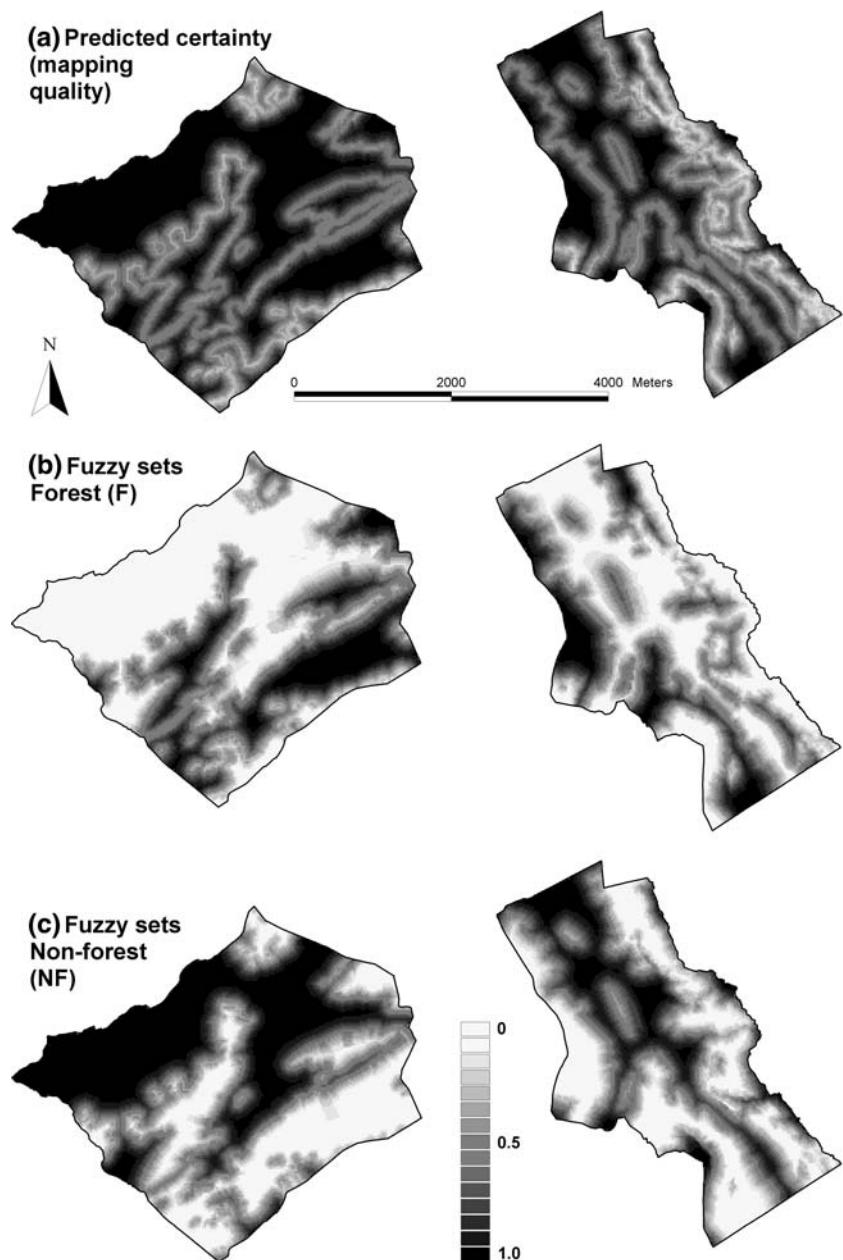
Bias correction, accuracy assessment and area estimation

Using the decision rules outlined in section 3.4, we corrected the forest area as mapped in the Siegfried Maps using NMI-based optimized thresholds. The new bias-corrected forest/non-forest maps showed higher accuracies than the original Siegfried Maps when tested against the reference maps of the respective other test area (Fig. 8; Table 2). The gain in accuracy for Pontresina is considerably higher (NMI changes from 0.26 to 0.38) than for St. Moritz (NMI changes from 0.36 to 0.45). Altogether, the proportion of

Table 1 Results from predictive uncertainty modelling. The models were calibrated in each area, and tested crosswise in the respective other area for independent accuracy assessment

	Calibration: St. Moritz Testing: Pontresina	Calibration: Pontresina Testing: St. Moritz
Model quality		
D^2	0.35	0.32
<i>Adjusted</i> D^2	0.35	0.31
Cross-wise Model test		
<i>Spearman</i> (ρ)	0.67	0.69
<i>G-value</i>	0.38	0.44
MAE	0.19	0.16
Model parameters		
Constant	0.3187	0.3073
Distance to forest boundary	0.01278	0.014
Elevation difference	– 0.02299	– 0.008461
Slope	– 0.001755	– 0.003259

Fig. 7 Cartographic distribution of fuzzy information in the two study areas St. Moritz (left) and Pontresina (right). The sub-panels show: **(a)** certainty values between 0 and 1 as predictive model output; fuzzy set values between 0 and 1 for **(b)** forest and **(c)** non-forest that are derived from predicted certainty values according to Eqs. 4 and 5



misclassified area was decreased from 6418 to 4873 pixels compared to the original map. The range of THR, for which the gain in classification accuracy reached its maximum in both study areas simultaneously, is between 0.49 and 0.57 (Fig. 8). This range indicates similar effects for different THR. The presented results (Table 2) and the improved maps (Fig. 9) were computed based on $\text{THR} = 0.54$.

An example of a land-cover change analysis

Forest cover change analysis was applied to both study areas using the forest cover of the modern map (Table 3). The deviation of the computed changes was decreased from 4960 units to 2602 units. Thus the procedure resulted in a reduction of misclassified land cover change units by nearly 50 percent. This improvement becomes especially

Fig. 8 Gain in map accuracy after reclassifying the fuzzy sets compared to the original map accuracy (NMI). The range of THR with high improvement is marked by the grey box

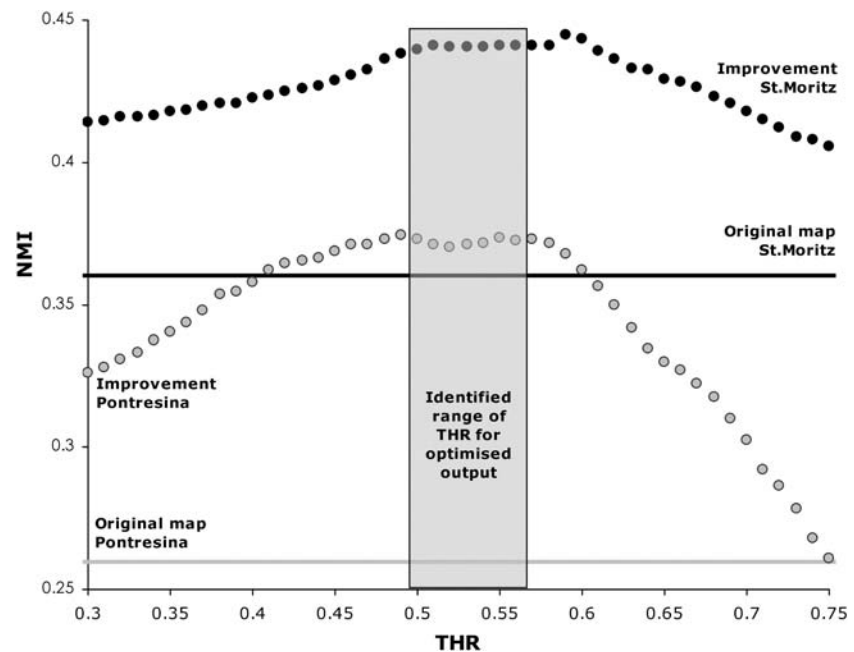


Table 2 Improvement of the historical map indicated by global accuracy measures and the decrease of misclassified proportions. The results are computed for $THR = 0.54$

	Reference map		Original map		Improved map	
	Pontresina	St. Moritz	Pontresina	St. Moritz	Pontresina	St. Moritz
Forest area (correct)	8350	8223	5853	6619	7080	7298
Non-forest area (correct)	6486	11496	5649	10016	5469	9835
Misclassified proportion	–	–	3334	3084	2287	2586
PCC	–	–	0.76	0.84	0.85	0.87
Kappa	–	–	0.55	0.67	0.69	0.74
NMI	–	–	0.26	0.36	0.38	0.45
THR value	0.54					

Fig. 9 Bias-corrected new crisp maps with the new forest units in black. The initial underestimation of forest area could be corrected, especially in regions of high uncertainty

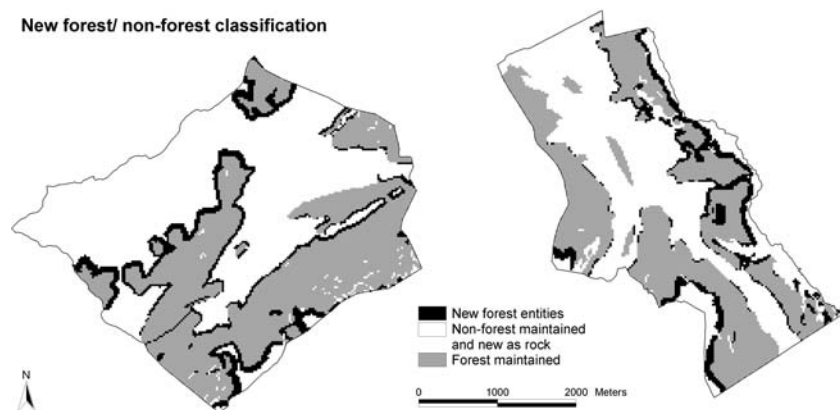


Table 3 Results from land-cover change analysis over both study areas and its improvement through bias correction in the historical map

Changes based on ...	Reference map		Original Siegfried Map		Improved Siegfried Map	
	Pontresina	St. Moritz	Pontresina	St. Moritz	Pontresina	St. Moritz
Non-forest to forest	614	772	1709	1322	959	742
Forest to non-forest	2341	1825	1912	2231	2420	2702
Forest maintained	5797	6398	4702	5785	5452	6365
Non-forest maintained	6084	10724	6513	10381	6005	9910
Deviations from reference analysis	–	–	3048	1912	848	1754

obvious in the Pontresina study area where this deviation was decreased from 3048 to 848 units (Table 3). In the St. Moritz study area forest area in the new historical map and thus the estimation of changes from forest to non-forest were slightly overestimated. Thus the improvement is less prominent.

Discussion

The described fuzzy set based procedure leads to a remarkable improvement of the original historical maps through the removal of classification bias. The approach allows us to confirm the hypothesised direction of the bias, i.e., forest was found to be underestimated in the Siegfried Maps. One reason for the underestimation could be that forest boundaries have been mapped from remote stand points (e.g., nearest road) if the terrain was not accessible and possibly difficult to overview. Thus forest memberships had to be weighted during the reclassification process using threshold-based decision rules. This resulted in an increase of forest area and independently tested accuracy. This gain in accuracy allowed more precise area estimations of the fuzzy classes since different meanings of forest and non-forest memberships were taken into account. This aspect has not been considered yet in recent research where the area of a fuzzy class was considered as a function of the alpha cut level only (Woodcock and Gopal 2000; Fonte and Lodwick 2004). The optimal values of THR for reclassifying the fuzzy maps were found to be similar in both test areas and their average provided a strong improvement in both maps (Table 2). If this similarity held for additional study areas, a general applicability of the approach to

other regions of similar topographic characteristics could be expected. Thus relationships between uncertainty and topographic predictors (Steele et al. 1998) were shown to be useful for predictive uncertainty modelling and for predictive, spatially explicit bias correction.

The model performance depends on the completeness of forest delineation in the investigated historical map. In general this completeness was satisfying in the Siegfried Map. Still, we noticed cases where forest gaps (e.g., avalanche tracks) or small forest patches were not delineated in the Siegfried Map. This is a limitation of the approach. Furthermore, we need to consider that there are various reasons for uncertainty in maps that cannot be addressed by such a model. Mainly attribute uncertainty remains unknown.

As could be seen in the example of St. Moritz the classification accuracy can only be improved to a certain degree. Here, the map was less distorted and forest cover was mapped with a higher accuracy ($\kappa = 0.67$) since the terrain is easier to access than in Pontresina ($\kappa = 0.55$). Topographers could move through this area more easily and better oversee the whole terrain. Thus, the bias is lower, and the gain in accuracy is also lower than in Pontresina. As a result, forest area was even slightly overestimated in some sub-areas of St. Moritz (Table 2).

The application example demonstrated that the procedure allowed for a more accurate forest cover change analysis compared to a reference change (Table 3). The quantitative information provided by the fuzzy sets could thus be directly used for change analyses by evaluating different thresholds used for reclassification. For more sophisticated change analyses the differences in forest definitions that are used in the maps to be compared, have to

be evaluated more carefully. Thus the semantic meaning of forest in the historical maps needs to be investigated to examine the degree of compatibility with definitions in modern maps. Such investigations will need to be linked to historical research since no scientific definitions of forest per se existed in the 19th century.

Conclusions

This paper contributes to the fields of uncertainty modelling, fuzzy set based spatial operations and land cover change analysis. It presents a model-based approach for correcting inherent uncertainty in land cover maps due to classification bias arising from underestimation of one class. In this approach we linked predictive uncertainty modelling using GLMs to fuzzy set based procedures. After identifying the direction of bias first, bias was then corrected by defining decision rules, which reflected different identification likelihoods during the historical mapping. This was done by pixel-wise weighting of forest memberships depending on the uncertainty level at that position. The result was a new crisp map of higher classification accuracy, which also allowed us to perform improved land cover change analyses. The method has been applied to historical forest/non-forest maps, which were derived from the Siegfried Map. Assuming that the results can be confirmed by more study areas, such a procedure could be applicable to other areas of similar topographic characteristics. Thus it overcomes the limitation of recent fuzzy set based approaches where classifications could be evaluated within areas covered by reference maps only (e.g. Binaghi et al. 1999; Jäger and Benz 2000).

The consideration of different meanings of forest and non-forest memberships by weighting them during the reclassification allowed us to compute more accurate area estimates of the fuzzy classes. Thus the approach proposes a further development of estimating the area of individual fuzzy classes as a function of the alpha cut level only (Woodcock and Gopal 2000; Fonte and Lodwick 2004). If more than one class is evaluated at the same time, the consideration of dif-

ferences in class-specific meanings results in more reliable area estimates. This is especially true if the classes have different transition characteristics due to varying amounts of inherent uncertainty, which can be related to different fuzzy object models presented by Cheng et al. (2001).

The procedure makes full use of the quantitative information provided by fuzzy sets. Thus subjectivity, which is inherent in fuzzy sets based on linguistic scales (Power et al. 2001; Woodcock and Gopal 2000), can be circumvented. The gain in accuracy is measurable and the predictive power of the uncertainty model can be tested for different study areas. Thus the approach is suitable for deriving more reliable knowledge of former landscape patterns at larger spatial domains of similar topographic characteristics. Thus it can be used for similar cases of land cover change models. However, it requires detailed and accurate reference maps in calibration areas that allow to train the uncertainty models.

Future research should be dedicated to the development of additional predictors for uncertainty modelling tasks applied to historical maps and alternative decision rules to produce a new corrected historical map based on fuzzy sets. For example, more than two certainty levels could be considered for defining thresholds. The incorporation of more than just two land cover classes would represent an interesting extension of the presented approach. Also, the direct use of weighted fuzzy set memberships for land cover change detection needs to be further investigated. The computation of areas of fuzzy classes, whose memberships have different meanings, needs to be further examined for other application examples.

Acknowledgements This study is part of the project “Retrospective Assessments” funded by the Swiss Agency for the Environment, Forests and Landscape and the Swiss Federal Research Institute for Forest, Snow, and Landscape. We would also like to thank the two reviewers for their valuable comments.

References

- Ahlqvist O, Keukelaar J, Oukbir K (2003) Rough and fuzzy geographical data integration. *Int J Geograph Inform Sci* 17(3):223–234

- Andréfouët S, Roux L, Chancerelle Y, Bonneville A (2000) A fuzzy-possibilistic scheme of study for objects with indeterminate boundaries: Application to French Polynesian reefscapes. *IEEE Trans Geosci Remote Sensing* 38(1):257–270
- Baker WL (1989) A review of models in landscape change. *Landscape Ecol* 2(2):111–133
- Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York
- Binaghi E, Brivio PA, Ghezzi P, Rampini A (1999) A fuzzy set-based accuracy assessment of soft classification. *Pattern Recognition Lett* 20:935–948
- Bolliger J, Mladenoff DJ (2005) Quantifying spatial classification uncertainties of the historical Wisconsin landscape (USA). *Ecography* 28:141–156
- Brown DG (1998) Classification and boundary vagueness in mapping pre-settlement forest types. *Int J Geograph Inform Sci* 12(2):105–129
- Burrough PA (1989) Fuzzy mathematical methods for soil survey and land evaluation. *J Soil Sci* 40:477–492
- Burrough P, McDonnell R (1998) Principles of geographical information systems. Oxford University Press
- Cheng T, Molenaar M, Lin H (2001) Formalizing fuzzy objects from uncertain classification results. *Int J Geograph Inform Sci* 15(1):27–42
- Cohen J (1960) A coefficient of agreement for nominal scales. *Edu Psychol Measurement* 20:37–46
- Coppin P, Jonckheere I, Nackaerts K, Muys B (2004) Digital change detection methods in ecosystem monitoring: a review. *Int J Remote Sensing* 25(9):1565–1596
- Dubois D, Prade H (2000) Fundamentals of fuzzy sets. The handbook of fuzzy sets series. Kluwer Academic, Dordrecht, The Netherlands
- Fisher P (2000) Sorites paradox and Vague Geographies. *Fuzzy Sets Systems* 113(1):7–18
- Fonte CC, Lodwick WA (2004) Areas of fuzzy geographical entities. *Int J Geograph Inform Sci* 18(2):127–150
- Foody GM (1996) Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *International Journal of Remote Sensing* 17(7):1317–1340
- Forbes AD (1995) Classification algorithm evaluation: five performance measures based on confusion matrices. *J Clin Monitoring* 11:189–206
- Gopal S, Woodcock C (1994) Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Eng Remote Sensing* 60(2):181–188
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Modelling* 135:147–186
- Insightful (2001) S-Plus 6 for Windows: User's Guide. Seattle7 Insightful, 688 pp
- Jäger G, Benz U (2000) Measures of classification accuracy based on fuzzy similarity. *IEEE Trans GeoSci Remote Sensing* 38:1462–1467
- Kienast F (1993) Analysis of historic landscape patterns with a Geographical Information system – a methodological outline. *Landscape Ecol* 8(2):103–118
- Klir GJ, Wierman MJ (1999) Uncertainty – based information – elements of generalized information theory. Springer, Physica-Verlag
- Krishnapuram R, Keller JM (1993) A possibilistic approach to clustering. *IEEE Trans Fuzzy Syst* 1:98–110
- Lewis HG, Brown M (2001) A generalised confusion matrix for assessing area estimates from remotely sensed data. *Int J Remote Sensing* 22(16):3223–3235
- Leyk S, Boesch R, Weibel R (2005) A conceptual framework for uncertainty investigation in map-based land cover change modelling. *Trans GIS* 9(3):291–322
- Leyk S, Zimmermann NE (2004) A predictive uncertainty model for field-based survey maps using Generalized Linear Models. In: Egenhofer M, Freksa C, Miller H (eds) *GIScience 2004. Lecture Notes in Computer Science* 3234. Springer, pp 191–205
- Lu D, Mausel P, Brondizio E, Moran E (2004) Change detection techniques. *Int J Remote Sensing* 25(12):2365–2407
- Matsakis P, Andréfouët S, Capolsini P (2000) Evaluation of fuzzy partitions. *Remote Sensing Environ* 74:516–533
- Plewe B (2002) The nature of uncertainty in historical geographic information. *Trans GIS* 6(4):431–456
- Power C, Simms A, White R (2001) Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *Int J Geograph Information Sci* 15(1):77–100
- Robinson VB (1988) Some implications of fuzzy set theory applied to geographic databases. *Computers, Environ Urban Systems* 12:89–97
- Robinson VB (2003) A perspective on the fundamentals of fuzzy sets and their use in Geographical Information Systems. *Trans GIS* 7(1):3–30
- Ruspini EH (1969) A new approach to clustering. *Information Control* 15:22–23
- Steele BM, Winne JC, Redmond RL (1998) Estimation and mapping of misclassification probabilities for thematic land cover maps. *Remote Sensing Environ* 66:192–202
- Woodcock CE, Gopal S (2000) Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *Int J Geograph Information Sci* 14:153–172
- Zadeh LA (1965) Fuzzy sets. *Information Control* 8:338–53